

Introduction to text mining Twitter with R

Paul Nulty

April 17, 2014

List of Acronyms and Definitions

- API: Application Programming Interface — a specification of how computer programs can communicate with each other
- JSON: JavaScript Object Notation — a human readable format for representing attribute-value pairs.
- OAuth: An open standard for authorization. OAuth provides client applications a ‘secure delegated access’ to server resources on behalf of a resource owner.
- HTTP: HyperText Transfer Protocol:
- REST: REpresentational State Transfer:

Useful Links

<https://dev.twitter.com/docs/things-every-developer-should-know>

```
2 + 3

## [1] 5

library(quanteda)
tokenize("Just testing sweave/knitr code integration here.")

## [1] "just"          "testing"        "sweaveknitr" "code"
## [5] "integration" "here"
```

The twitter APIs

Twitter offers two ways for computer programs to post and retrieve data from its service. The twitter REST Search API will receive requests as http commands, and will respond with a JSON object representing a list of tweets. The twitter streaming API allows a computer program to maintain a connection to twitter's service and receive a flow of live public tweets that match the filters or search terms specified.

Both of these APIs require that the developer of the computer program create a twitter account, create a new application, and receive an access token.

Getting programmatic access to twitter

1. Go to <https://dev.twitter.com/> and sign in with a twitter account
2. Go to My Applications and select "New App". <https://apps.twitter.com/>
3. Fill out the form and click 'create'